

# Lepton flavour violating Higgs boson decay $h \rightarrow \mu\tau$ at the ILC

Indrani Chakraborty <sup>\*</sup>, Amitava Datta <sup>†</sup>, and Anirban Kundu <sup>‡</sup>

*Department of Physics, University of Calcutta,  
92 Acharya Prafulla Chandra Road, Kolkata 700009, India*

## Abstract

We study the possible reach of the proposed International Linear Collider (ILC) in exploring the lepton flavour violating (LFV) Higgs boson decay  $h \rightarrow \mu\tau$ . Two generic types of models are investigated, both involving an extended scalar sector. For the first type, the rest of the scalars are heavy and beyond the reach of ILC but the LFV decay occurs through a tiny admixture of the Standard Model (SM) doublet with the heavy degrees of freedom. In the second class, which is more constrained from the existing data, the SM Higgs boson does not have any LFV decay but there are other scalars degenerate with it that gives rise to the LFV signal. We show that ILC can pin down the branching fraction of  $h \rightarrow \mu\tau$ , and hence the effective LFV Yukawa coupling, to a very small value, due to the fact that there are signal channels with unlike flavour leptons but no missing energy. It turns out that the low-energy options of the ILC, namely,  $\sqrt{s} = 250$  or 500 GeV are better to investigate such channels, and the option of beam polarization helps too. At least an order of magnitude improvement is envisaged over the existing limits, and the effective LFV Yukawa coupling can be probed at the level of  $10^{-4}$ .

PACS no.: 12.60.Fr, 13.66.Fg

## 1 Introduction

The leptonic flavour violating (LFV) decay of the Higgs boson  $h \rightarrow \mu\tau$ , which is a smoking gun for physics beyond the Standard Model (SM), has been looked for by both CMS and ATLAS collaborations at the Large Hadron Collider (LHC), and the branching ratio (BR) is [1, 2]

$$\text{BR}(h \rightarrow \mu\tau) = 0.84^{+0.39}_{-0.37}\% \text{ (CMS)}, \quad 0.53 \pm 0.51\% \text{ (ATLAS)}, \quad (1)$$

so that the 95% CL upper limits on the BR is 1.51% (CMS) and 1.41% (ATLAS). While there is no official averaging, one might take  $\text{BR}(h \rightarrow \mu\tau) < 1.4\%$ .

This has given rise to a lot of model building efforts to explain the signal [3, 4, 5, 6, 7]. As the  $\mu\tau$  pair has an invariant mass suggestive of a parent of mass close to that of the Higgs boson, there are only two possibilities: (i) By some mechanism (possibly from mixing with other scalars) the SM Higgs boson has acquired some small flavour-changing (FC) coupling; (ii) The leptons are coming from some other scalar degenerate with the Higgs boson. Both possibilities have been explored in the literature [4, 5, 6, 7], although the second possibility appears to be more constrained from the data. There is also a third possibility of introducing the LFV coupling through loop effects without enhancing the SM scalar sector (but other new degrees of freedom may be needed) [8]; however, this is equivalent to the introduction of an effective LFV coupling, as is the case for the first possibility.

Let us assume that the decay occurs through a term  $\left[-y_{ij}\bar{\ell}_L^i \ell_R^j h + \text{h.c.}\right]$  in the Lagrangian, where  $h$  is the scalar resonance at 125 GeV. Written in full, the relevant term is

$$-y_{\mu\tau}(\bar{\mu}_L\tau_R + \bar{\tau}_R\mu_L)h - y_{\tau\mu}(\bar{\tau}_L\mu_R + \bar{\mu}_R\tau_L)h. \quad (2)$$

The BR is given by

$$\text{BR}(h \rightarrow \mu\tau) = \frac{m_h}{8\pi\Gamma_h}(|y_{\tau\mu}|^2 + |y_{\mu\tau}|^2) \equiv \frac{m_h}{8\pi\Gamma_h}\mathcal{F}^2. \quad (3)$$

<sup>\*</sup>indrani300888@gmail.com

<sup>†</sup>adatta\_ju@yahoo.com

<sup>‡</sup>anirban.kundu.cu@gmail.com

If  $h \rightarrow \mu\tau$  (and other possible new decay channels) have a small BR, one can write  $\Gamma_h \approx \Gamma_h^{\text{SM}} = 4.07 \text{ MeV}$  for  $m_h = 125 \text{ GeV}$ .

The measurement constrains  $\mathcal{F} = \sqrt{|y_{\mu\tau}|^2 + |y_{\tau\mu}|^2}$ . Unless we measure the  $\mu$  or  $\tau$  polarizations, no further information about the individual couplings can be drawn. The upper bound from ATLAS [2] suggests

$$\mathcal{F} < 3.4 \times 10^{-3} \quad (4)$$

at 95% CL. If there is a hierarchy between the two couplings  $y_{\mu\tau}$  and  $y_{\tau\mu}$ , this is the bound on the larger coupling. If both are equal, the bound on each of them is  $1/\sqrt{2}$  times that on  $\mathcal{F}$ . We will show how far  $\mathcal{F}$  can be probed. Its relationship with the Yukawa couplings and mixing angles is model-dependent.

If the new channels have a significant BR, one should use

$$\Gamma_h = \frac{\Gamma_h^{\text{SM}}}{1 - \text{BR}(h \rightarrow \text{new})} \quad (5)$$

Note that such a scenario is quite feasible. Experimentally,  $\Gamma_h$  has a strong lower bound but a very weak upper bound, so even larger values of  $\mathcal{F}$  can be allowed in principle. The upper limit on  $\Gamma_h$  from direct measurement is  $1.7 \text{ GeV}$  [9] and from indirect measurement,  $22 \text{ MeV}$  [10]. This makes the  $1\sigma$  upper bound on  $\mathcal{F}$  to be  $0.063$  (direct measurement) or  $0.007$  (indirect measurement) approximately, assuming the only new channel to be  $h \rightarrow \mu\tau$ , for which one may allow a large branching ratio.

We would like to find out how far the FC coupling(s) can be explored at the International Linear Collider (ILC) [11]. While there are such studies in the literature in the context of supersymmetric LFV models [12], we would like to give more general constraints based on the updated ILC parameters. Almost all the models that try to explain the FC decay invoke an extended scalar sector<sup>1</sup>. As we have said before, there are generically two classes of models. The first class, which we may call the Lonely Higgs (LH) models, has an extended scalar sector, but all the scalars except the Higgs boson  $h$  are heavy. One or more of these extra degrees of freedom may mix with the SM doublet, and that gives rise to a small FC coupling of  $h$ . If the FC Yukawa coupling of the heavy scalars be  $y_H$  and the mixing angle with the SM doublet be  $\theta$ , we expect  $\mathcal{F} \sim y_H \sin \theta$ . The precise relation between  $\mathcal{F}$  and  $\{y_H, \theta\}$  depends on the model chosen and its parameters. The LH models have been widely investigated in the literature and include Type-III 2HDM [4], or scalar sectors with discrete symmetries, or even composite Higgs models. For example, in Ref. [4], it has been shown that in a flavour-changing 2HDM, the effective Yukawa coupling of  $\bar{\tau}\mu h$  is a function of the angles  $\alpha$  (the mixing angle of the two CP-even neutral states),  $\beta$  (the mixing angle of the charged scalar states, and the ratio of the two vacuum expectation values  $v_2/v_1$  is given by  $\tan \beta$ ), and  $\theta_R$  (the mixing angle of the charged leptons  $\mu_R$  and  $\tau_R$ ), apart from the FC Yukawa coupling of the second doublet. Since the functional form depends on the model chosen, we will not go into any particular model and express all our results in terms of  $\mathcal{F}$  as defined before. We will also assume all other scalars of such LH models to be beyond the reach of ILC.

The second class, which we may call the Degenerate Higgs (DH) models, involves, apart from  $h$ , more scalars at  $125 \text{ GeV}$ , some (or all) of them having nonzero FC couplings and giving rise to the  $\mu$ - $\tau$  signal. Examples of DH models are in Refs. [5, 13]. As these models are not very well investigated but may be interesting from an experimental point of view regarding clear signals or possible falsifiability, let us give a very brief outline of one of the DH models discussed in [5].

In this model, all the neutral scalars are almost degenerate. The SM Higgs boson  $h$  has only flavour-conserving couplings; the other CP-even neutral scalar  $H$  and the CP-odd pseudoscalar  $A$  can decay to the flavour-violating channels. The approximate degeneracy is forced by the observed signal, where the  $\mu$ - $\tau$  invariant mass peaks at  $m_h$ .

In the exact alignment limit, with  $\alpha$  and  $\beta$  defined before,

$$|\alpha - \beta| = (2n + 1)\frac{\pi}{2}, \quad (6)$$

---

<sup>1</sup>LFV couplings of the Higgs boson must identically vanish if the mass matrix is proportional to the Yukawa matrix, as the off-diagonal Yukawa interactions — even if effective — vanish in the stationary basis. With a single Higgs doublet, one needs to invoke higher-dimensional operators for this that can spoil the proportionality between the mass and the Yukawa matrices.

which makes the lighter CP-even eigenstate identical in properties with the SM Higgs boson<sup>2</sup>. This is the limit that is favored by the LHC data, but small deviations can still be entertained. The amount of the allowed deviation depends on the type of 2HDM [14].

The pseudoscalar  $A$  does not have any  $WWA$  or  $ZZA$  coupling, so it can never be produced by the Bjorken process or gauge boson fusion<sup>3</sup>. If one is away from the alignment limit, the  $h$  production rate is  $\sigma(h_{SM}) \times \sin^2(\beta - \alpha)$ , and the  $H$  production rate is  $\sigma(h_{SM}) \times \cos^2(\beta - \alpha)$ , where  $\sigma(h_{SM})$  is the SM Higgs boson production cross-section. On the other hand, the FC coupling of  $h$  ( $H$ ) is proportional to  $\cos(\sin)(\beta - \alpha)$ . Thus,  $\mathcal{F} \sim y \sin 2(\beta - \alpha)$  where  $y$  is the (effective) FC Yukawa coupling of  $H$ , like  $y_{\mu\tau}$  or  $y_{\tau\mu}$  as shown in Eq. (2). One must mention that the simplest version of this model as in Ref. [5] does not work because the model predicts a large branching ratio for  $h \rightarrow e\tau$  which has not been observed by the CMS collaboration [15]. Still, one may tweak the model and have a deviation from the strict alignment limit to satisfy all the experimental constraints [7].

Thus, what one can bound from the data is either (i) the Yukawa coupling of the  $\mu$ - $\tau$  pair with the heavy scalar, multiplied by  $\sin\theta$ , where  $\theta$  is the mixing angle between the SM Higgs and the heavy neutral one, for the LH models, or (ii) the Yukawa coupling of the other neutral scalar  $H$  to  $\mu\tau$  multiplied by a measure of the deviation from the alignment limit, for the DH models. Any further information would require a knowledge of the model, as investigated in Ref. [6]. Various LFV models were also studied in the literature before this signal came into prominence; for example, we refer the reader to Ref. [16].

## 2 The ILC parameters

The designed CM energies for the ILC are 250 and 500 GeV. While it can be upgraded to 1 TeV in future, we will work with the two designed energies only. The reason is that the sensitivity to LFV couplings are best at the low energy options. ILC is planned to run with four polarization options: P1  $(-0.8, 0.3)$ , P2  $(0.8, -0.3)$ , P3  $(-0.8, -0.3)$  and P4  $(0.8, 0.3)$  respectively [17, 18, 19]. The numbers give the degree of polarization of the beams; for example, the first option (P1) means that the  $e^-$  beam is 80% left polarized and the  $e^+$  beam is 30% right polarized. The effective polarization is given by

$$P_{\text{eff}} = \frac{P(e^-) + P(e^+)}{1 + P(e^-)P(e^+)}, \quad (7)$$

where  $P(e^-)$  is positive for left polarized electron beam and  $P(e^+)$  is positive for right polarized positron beam. The projected integrated luminosities for the four polarization options at  $\sqrt{s} = 250$  GeV, are (in  $\text{fb}^{-1}$ ):

$$\text{P1 : } 337.5, \quad \text{P2 : } 112.5, \quad \text{P3 : } 25, \quad \text{P4 : } 25, \quad (8)$$

which will finally be upgraded to 1350, 450, 100, 100  $\text{fb}^{-1}$  respectively. At  $\sqrt{s} = 500$  GeV, initial (upgraded) integrated luminosities are (in  $\text{fb}^{-1}$ ):

$$\text{P1 : } 200(1600), \quad \text{P2 : } 200(1600), \quad \text{P3 : } 50(400), \quad \text{P4 : } 50(400). \quad (9)$$

## 3 Signal and background

The dominant Higgs production processes at the ILC are the Bjorken process  $e^+e^- \rightarrow Zh$  and the vector boson fusion (VBF)  $e^+e^- \rightarrow h\nu_e\bar{\nu}_e(h e^+e^-)$  through  $WW(ZZ)$  fusion. The cross-section for the former falls with  $\sqrt{s}$  and that of the latter rises. At  $\sqrt{s} = 250$  GeV, the Bjorken process dominates, while at  $\sqrt{s} = 500$  GeV, both the processes become compatible.

The final state that we look for is  $q\bar{q}\mu\tau$  or  $\ell^+\ell^-\mu\tau$ , with  $\ell = e, \mu$ . To reliably reconstruct the  $\tau$  from its one-prong or three-prong decays, one should use the collinear approximation, *i.e.*, the  $\tau$  is so boosted that the

<sup>2</sup> In the other alignment limit  $|\alpha - \beta| = n\pi$ , the heavy CP-even eigenstate is identical to the SM Higgs. This possibility is also not ruled out from the LHC data.

<sup>3</sup> At an integrated luminosity of 500  $\text{fb}^{-1}$ , double scalar productions from the two gauge-two scalar quartic interactions are anyway negligibly small.

neutrino goes essentially in the same direction as its parent. If there are no other sources of missing energy, this reconstruction is quite reliable [17]. Thus, we will not consider the  $WW$  fusion further, first because the contributions are less than those coming from the Bjorken process, and second, because the extra neutrinos make the  $\tau$  reconstruction more uncertain. It is unlikely that to get the LFV signal from a 125 GeV resonance, one has to wait till the 1 TeV upgrade, when the initial options, through the Bjorken process, provide a better sensitivity<sup>4</sup>. If we focus only on the Bjorken process, there are no further sources of missing energy and the  $\tau$  reconstruction efficiency  $\epsilon_\tau$  is about 70% [17]. In the clean environment of the  $e^+e^-$  collider, the two jets can be detected with almost 100% efficiency (we do not need to tag the flavour), and so can be the muon. The electron detection efficiency is about 89%, while that of the muon is almost 100%. Note that the  $\mu\mu\mu\tau$  channel involves an extra combinatoric, *i.e.* one needs to check which muon pair came from the  $Z$  and which one came from the Higgs boson. This reduces the efficiency of the three-muon channel by about 20% compared to the  $ee\mu\tau$  channel, which approximately offsets the gain coming from the better detection efficiency. A detailed discussion will follow.

There is, of course, one possible source of background as discussed in Ref. [12]. This comes from the decay  $h \rightarrow \tau^+\tau^-$ , with one of the tau leptons decaying into a muon. If the missing energy is small, this can fake the signal event. We have calculated the number of such fake events,  $N_{\text{fake}}$ , for different ILC configurations and final states. An explicit evaluation (at parton level) was performed by CalcHEP, generating the  $q\bar{q}(e^+e^-)\mu\tau\nu_\mu\nu_\tau$  final state at the ILC, including all possible amplitudes. We then applied the cuts as mentioned later: the  $q\bar{q}(e^+e^-)$  invariant mass at an interval of  $m_Z \pm 10$  GeV, the  $\mu\tau$  invariant mass within  $m_h \pm 10$  GeV, and missing energy less than 20 GeV. The cuts reduce the uncut cross-section by at least 3 orders of magnitude. This is expected as the  $\mu\tau$  pair can come from many sources, not necessarily from  $h \rightarrow \tau\tau$ . After multiplication by the  $\tau$  detection efficiency  $\epsilon_\tau = 0.7$  and the projected luminosity, the number of fake events becomes completely negligible. Slightly conservative estimate is given in Ref. [12] which we show in the tables. While the number of such fake events  $N_{\text{fake}}$  is completely negligible for  $e^+e^-\mu\tau$  final states, it is quite small (and even zero for some configurations) for  $q\bar{q}\mu\tau$  final states too. One of the reasons for this is that the number of signal as well as parent background (with  $h \rightarrow \tau^+\tau^-$ ) events are much smaller for the former case compared to the latter, coming from  $\text{BR}(Z \rightarrow e^+e^-)/\text{BR}(Z \rightarrow q\bar{q})$ . The second reason is that the tau lepton can be more easily reconstructed in the former case than the latter.

The cross-section for the Bjorken process and the subdominant  $ZZ$  fusion are shown in Table 1.

$\sqrt{s}$ (GeV)	Polarization	Cross-section (fb)	
		Bjorken	$ZZ$ fusion
250	P1	333	0.81
	P2	227	0.52
	P3	196	0.75
	P4	148	0.62
500	P1	105	8.62
	P2	71.4	5.48
	P3	61.6	7.66
	P4	46.4	6.25

Table 1: Higgs production cross-section for different  $\sqrt{s}$  and polarization options at the ILC. Computed with CalcHEP v3.6.23 [20] with ISR and beamstrahlung options switched on.

We will consider the following signal processes.

1.  $e^+e^- \rightarrow Zh$ ,  $Z \rightarrow q\bar{q}$ ,  $h \rightarrow \mu\tau$ : This is the most efficient channel to probe the FC coupling(s) of the SM Higgs boson in the early (low- $\sqrt{s}$ ) runs of the ILC. The channel has almost zero SM background if we use the cuts mentioned above — we take the invariant mass of the two jets in the interval  $m_Z \pm 10$  GeV, the invariant mass of the  $\mu\tau$  pair in the interval  $m_h \pm 10$  GeV, and missing energy below 20 GeV. With these

<sup>4</sup>There might be heavier scalars with LFV couplings too, and for them the high- $\sqrt{s}$  option is a must, either at the ILC or the  $e^+e^-$  mode of the Future Circular Collider (FCC-ee).

cuts, one retains 90% or more of the signal events while the backgrounds coming from  $t\bar{t}$  or  $W^+W^-Z/\gamma$  final states can be successfully eliminated<sup>5</sup>. Note that any SM process with two different flavoured leptons must involve at least two neutrinos and therefore a significant missing energy, unless the neutrinos conspire otherwise.

2.  $e^+e^- \rightarrow Zh, Z \rightarrow \ell^+\ell^-, h \rightarrow \mu\tau$ , with  $\ell = e, \mu$ : The reach is significantly lower than  $Z \rightarrow q\bar{q}$  because of the BR suppression for  $Z \rightarrow \ell^+\ell^-$ . For  $Z \rightarrow \mu^+\mu^-$ , one has to apply combinatorics for muons in the final state to determine which of the two like-sign muons came from  $h$  and which from  $Z$ . Consider, for example, the final state  $\mu^+\mu^-\mu^+\tau^-$ . For the event to be acceptable, we have to know for sure which  $\mu^+$  originated from  $Z$  and this is not possible if the invariant masses of the  $\mu^-$  with both  $\mu^+$  fall close to the  $Z$ -peak, say in the  $m_Z \pm 5$  GeV range. We have checked that this amounts to a reduction of 20% at the most. At the same time, the detection efficiency of  $\mu$  is almost 100% compared to 89% for the electron; which means that one is likely to detect equal number of  $ee\mu\tau$  and  $\mu\mu\mu\tau$  events. There is a small contribution to  $ee\mu\tau$  final state from  $ZZ$  fusion which is absent for the three-muon final state. Again, the cuts remove all the backgrounds but never more than 10% of the signal events. We found  $N_{\text{fake}}$  to be always zero for this case.

Our results are shown in the following tables. We take the integrated luminosity as envisaged by the ILC, and the detection efficiencies as mentioned before:

$$\epsilon_\tau = 0.7, \quad \epsilon_\mu = \epsilon_{\text{jet}} = 1.0, \quad \epsilon_e = 0.89, \quad (10)$$

and the branching ratios

$$\text{BR}(Z \rightarrow q\bar{q}) = 0.7, \quad \text{BR}(Z \rightarrow \nu\bar{\nu}) = 0.2, \quad \text{BR}(Z \rightarrow \ell^+\ell^-) = 0.03363. \quad (11)$$

Table 2 shows the number of events for various  $\sqrt{s}$  and polarization options if  $\text{BR}(h \rightarrow \mu\tau) = 0.01$ , where we have taken into account the charge-conjugated final state too. The cuts as mentioned above have been applied. From the large number of events in channels with zero or small background, it is clear that ILC can probe to a far smaller BR.

	$\sqrt{s}$ (GeV)	Polarization	$Zh, Z \rightarrow q\bar{q}$	$Zh, Z \rightarrow e^+e^-$	$N_{\text{fake}}$
Initial	250	P1	496	23	2
		P2	113	5	0
		P3	22	1	0
		P4	16	1	0
Final	250	P1	1983	92	8
		P2	450	21	2
		P3	86	5	0
		P4	65	3	0
Initial	500	P1	93	14	0
		P2	63	10	0
		P3	13	3	0
		P4	10	3	0
Final	500	P1	741	119	3
		P2	504	77	2
		P3	109	24	0
		P4	82	19	0

Table 2: Number of events, after a reduction by 10% for the cuts applied, assuming  $\text{BR}(h \rightarrow \mu\tau) = 0.01$  for different  $\mathcal{L}$  and polarization options as mentioned in the text.  $N_{\text{fake}}$  denotes the number of fake events for  $q\bar{q}\mu\tau$  final states. There are no fake events at this luminosity in the  $ee(\mu\mu)\mu\tau$  channels. The number of  $\mu\mu\mu\tau$  events is approximately the same as the number of  $ee\mu\tau$  events because of a cancellation between better detection efficiency and combinatoric loss; see text for details.

<sup>5</sup> $t\bar{t}$  final states can give rise to two  $b$ -jets,  $\mu^\pm\tau^\mp$ , and missing energy. For  $W^+W^-Z/\gamma$ , one gets the same background (with missing energy) when the two  $W$ 's decay leptonically, and the  $Z/\gamma$  goes to  $q\bar{q}$ .

	$\sqrt{s}$ (GeV)	Polarization	$\mathcal{L}$ (fb $^{-1}$ )	BR(5, $\mathcal{L}$ )	$\mathcal{F}(5)$	BR(3, $\mathcal{L}$ )	$\mathcal{F}(3)$
Initial	250	P1	337.5	$1.1 \times 10^{-4}$	$2.9 \times 10^{-4}$	$7.3 \times 10^{-5}$	$2.4 \times 10^{-4}$
		P2	112.5	$3.9 \times 10^{-4}$	$5.7 \times 10^{-4}$	$2.4 \times 10^{-4}$	$4.4 \times 10^{-4}$
		P3	25	$2.1 \times 10^{-3}$	$1.3 \times 10^{-3}$	$1.2 \times 10^{-3}$	$1.0 \times 10^{-3}$
		P4	25	$2.7 \times 10^{-3}$	$1.5 \times 10^{-3}$	$1.6 \times 10^{-3}$	$1.2 \times 10^{-3}$
Final	250	P1	1350	$4.1 \times 10^{-5}$	$1.8 \times 10^{-4}$	$3.2 \times 10^{-5}$	$1.6 \times 10^{-4}$
		P2	450	$1.2 \times 10^{-4}$	$3.1 \times 10^{-4}$	$7.9 \times 10^{-5}$	$2.5 \times 10^{-4}$
		P3	100	$5.2 \times 10^{-4}$	$6.5 \times 10^{-4}$	$3.1 \times 10^{-4}$	$5.0 \times 10^{-4}$
		P4	100	$6.9 \times 10^{-4}$	$7.5 \times 10^{-4}$	$4.1 \times 10^{-4}$	$5.8 \times 10^{-4}$
Initial	500	P1	200	$4.8 \times 10^{-4}$	$6.3 \times 10^{-4}$	$2.9 \times 10^{-4}$	$4.9 \times 10^{-4}$
		P2	200	$7.1 \times 10^{-4}$	$7.6 \times 10^{-4}$	$4.3 \times 10^{-4}$	$5.9 \times 10^{-4}$
		P3	50	$3.3 \times 10^{-3}$	$1.6 \times 10^{-3}$	$1.9 \times 10^{-3}$	$1.3 \times 10^{-3}$
		P4	50	$4.4 \times 10^{-3}$	$1.9 \times 10^{-3}$	$2.6 \times 10^{-3}$	$1.5 \times 10^{-3}$
Final	500	P1	1600	$8.5 \times 10^{-5}$	$2.6 \times 10^{-4}$	$6.1 \times 10^{-5}$	$2.2 \times 10^{-4}$
		P2	1600	$1.1 \times 10^{-4}$	$2.9 \times 10^{-4}$	$7.1 \times 10^{-5}$	$2.4 \times 10^{-4}$
		P3	400	$4.1 \times 10^{-4}$	$5.8 \times 10^{-4}$	$2.5 \times 10^{-4}$	$4.5 \times 10^{-4}$
		P4	400	$5.5 \times 10^{-4}$	$6.7 \times 10^{-4}$	$3.3 \times 10^{-4}$	$5.2 \times 10^{-4}$

Table 3: Reach of  $\mathcal{F}$  in the  $\mu\tau q\bar{q}$  channel, for observing 5 events (columns 5 and 6) or 3 events (columns 7 and 8) at the designed  $\mathcal{L}$  and polarization options. This is based on either 5(3) events where the background is zero, or a significance of  $5\sigma(3\sigma)$  where the background is nonzero.

For brevity, let us denote the cross-sections for the Bjorken process and  $ZZ$  fusion by  $\sigma_b$  and  $\sigma_{ZZ}$  respectively. If  $N$  events are necessary to claim a discovery where the background is negligible, the BR can be probed up to

$$\text{BR}(N, \mathcal{L}) = \frac{N}{\sigma_b Z_f \epsilon_f \epsilon_\tau \epsilon_\mu \mathcal{L}}, \quad (12)$$

where  $Z_f$  is BR of  $Z$  in the corresponding fermion channel, as shown in Eq. (11), and  $\epsilon_f = \epsilon_{\text{jet}}, \epsilon_e$  and  $\epsilon_\mu$  for the  $q\bar{q}$ ,  $e^+e^-$ , and  $\mu^+\mu^-$  channels respectively.  $\mathcal{L}$  is the integrated luminosity in fb $^{-1}$ . We quote our numbers for  $N = 5$  and  $N = 3$  in Table 3. Also shown are the  $5\sigma$  and  $3\sigma$  ranges where the background coming from fake events are nonzero, using the standard  $S/\sqrt{S+B}$  measure for the significance. Note that only for the  $e^+e^-\mu\tau$  final states,  $\sigma_b Z_f \epsilon_f$  is to be replaced by  $\sigma_b Z_f \epsilon_f + \sigma_{ZZ}$ . For leptonic final states, we add both  $ee\mu\tau$  and  $\mu\mu\mu\tau$  events to increase the statistics.

The range of  $\mathcal{F}$  that can be probed is thus given by

$$\mathcal{F} = \sqrt{\frac{8\pi\Gamma_h \text{BR}(N, \mathcal{L})}{m_h}}. \quad (13)$$

The best results are obtained for  $\mu\tau q\bar{q}$  final states and for low- $\sqrt{s}$  options of ILC. As we see from Table 3, the maximum reach can be obtained for  $\sqrt{s} = 250$  GeV, and the polarization options also help. With zero background, even 5 signal events mean a discovery, or a  $5\sigma$  significance with nonzero background. With the cuts and detection efficiencies as mentioned, a  $5\sigma$  significance with the P1 option at  $\sqrt{s} = 250$  GeV and integrated luminosity  $\mathcal{L} = 1350$  fb $^{-1}$  gives  $\text{BR}(h \rightarrow \mu\tau) \approx 4.1 \times 10^{-5}$ , or  $\mathcal{F} \approx 1.8 \times 10^{-4}$ . This is the best reach of the ILC. Obviously, the reach will be somewhat better with lesser significance.

The reach in the  $\ell^+\ell^-\mu\tau$  channel is much worse compared to that of  $q\bar{q}\mu\tau$ , even though the former is background-free, because of the lower BR of  $Z$  to  $\ell^+\ell^-$ . The results are displayed in Table 4. A comparative idea of different channels can also be found in Fig. 1.

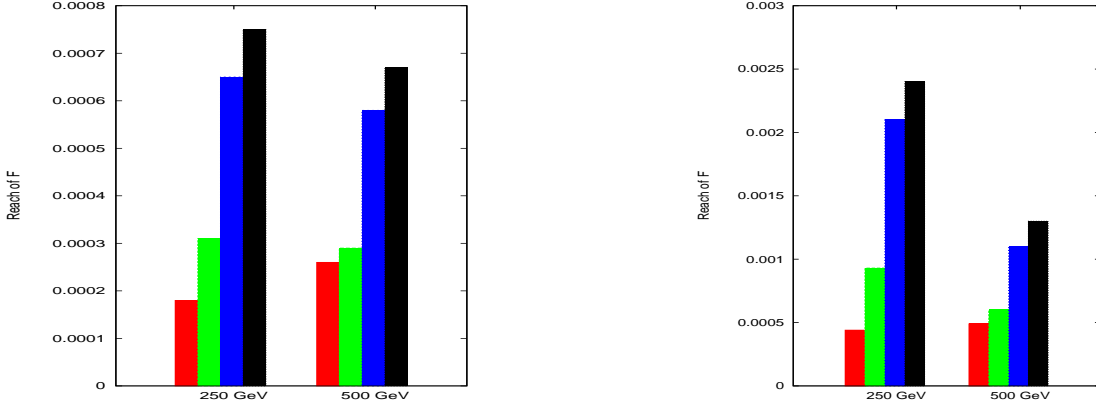


Figure 1: The reach for  $\mathcal{F}$  (see text for the definition) for different choices of  $\sqrt{s}$ , and with the larger (*i.e.*, final) luminosity option. The red, green, blue, and black bars are respectively for the polarization options P1, P2, P3 and P4 with four different integrated luminosities mentioned earlier. The left (right) plot shows the reach for the  $\mu\tau q\bar{q}$  ( $\mu\tau\ell^+\ell^-$ ) channel. With nonzero backgrounds, the bars show the  $5\sigma$  reach; with zero background, the reach with 5 events.

	$\sqrt{s}$ (GeV)	Polarization	$\mathcal{L}$ (fb $^{-1}$ )	BR(5, $\mathcal{L}$ )	$\mathcal{F}(5)$	BR(3, $\mathcal{L}$ )	$\mathcal{F}(3)$
Initial	250	P1	337.5	$9.6 \times 10^{-4}$	$8.9 \times 10^{-4}$	$5.8 \times 10^{-4}$	$6.9 \times 10^{-4}$
		P2	112.5	$4.2 \times 10^{-3}$	$1.9 \times 10^{-3}$	$2.5 \times 10^{-3}$	$1.4 \times 10^{-3}$
		P3	25	$2.2 \times 10^{-2}$	$4.2 \times 10^{-3}$	$1.3 \times 10^{-2}$	$3.3 \times 10^{-3}$
		P4	25	$2.8 \times 10^{-2}$	$4.9 \times 10^{-3}$	$1.7 \times 10^{-2}$	$3.8 \times 10^{-3}$
Final	250	P1	1350	$2.4 \times 10^{-4}$	$4.4 \times 10^{-4}$	$1.4 \times 10^{-4}$	$3.4 \times 10^{-4}$
		P2	450	$1.1 \times 10^{-3}$	$9.3 \times 10^{-4}$	$6.4 \times 10^{-4}$	$7.2 \times 10^{-4}$
		P3	100	$5.4 \times 10^{-3}$	$2.1 \times 10^{-3}$	$3.2 \times 10^{-3}$	$1.6 \times 10^{-3}$
		P4	100	$7.1 \times 10^{-3}$	$2.4 \times 10^{-3}$	$4.3 \times 10^{-3}$	$1.9 \times 10^{-3}$
Initial	500	P1	200	$2.3 \times 10^{-3}$	$1.4 \times 10^{-3}$	$1.4 \times 10^{-3}$	$1.1 \times 10^{-3}$
		P2	200	$3.6 \times 10^{-3}$	$1.7 \times 10^{-3}$	$2.1 \times 10^{-3}$	$1.3 \times 10^{-3}$
		P3	50	$1.2 \times 10^{-2}$	$3.2 \times 10^{-3}$	$7.4 \times 10^{-3}$	$2.5 \times 10^{-3}$
		P4	50	$1.5 \times 10^{-2}$	$3.6 \times 10^{-3}$	$9.3 \times 10^{-3}$	$2.8 \times 10^{-3}$
Final	500	P1	1600	$2.9 \times 10^{-4}$	$4.9 \times 10^{-4}$	$1.7 \times 10^{-4}$	$3.8 \times 10^{-4}$
		P2	1600	$4.4 \times 10^{-4}$	$6.0 \times 10^{-4}$	$2.7 \times 10^{-4}$	$4.7 \times 10^{-4}$
		P3	400	$1.5 \times 10^{-3}$	$1.1 \times 10^{-3}$	$9.2 \times 10^{-4}$	$8.7 \times 10^{-4}$
		P4	400	$1.9 \times 10^{-3}$	$1.3 \times 10^{-3}$	$1.2 \times 10^{-3}$	$9.8 \times 10^{-4}$

Table 4: Reach of  $\mathcal{F}$  in the  $\mu\tau\ell^+\ell^-$  channel, summing over  $\ell = e$  and  $\ell = \mu$ , for observing 5(3) events at the designed  $\mathcal{L}$  and polarization options. This channel has negligible  $N_{\text{fake}}$ .

## 4 Summary

In view of the LFV channel  $h \rightarrow \mu\tau$  with a substantially strong hint at the LHC, we investigate the reach of ILC to probe the corresponding LFV Yukawa coupling. The advantage of ILC is a relatively clean environment and much less QCD background compared to the LHC. An important aspect of the signal is that any SM background has to contain missing  $p_T$  coming from the two neutrinos. Thus, if the Higgs boson is produced through the Bjorken process and the  $Z$  decays to two jets or two leptons, one may have a signal which is background-free. This helps us to achieve at least one order of improvement to what we have at the LHC now. The reach for the Yukawa coupling  $\mathcal{F}$  also depends on the integrated luminosity  $\mathcal{L}$  and goes as  $1/\sqrt{\mathcal{L}}$ . With an integrated luminosity of  $\mathcal{L} = 1350 \text{ fb}^{-1}$ ,  $\text{BR}(h \rightarrow \mu\tau)$  can be probed up to a level of  $4.1 \times 10^{-5}$  even at  $\sqrt{s} = 250 \text{ GeV}$ , which in turn means  $\mathcal{F} \approx 1.8 \times 10^{-4}$ . This is more than an order of magnitude improvement over the LHC. The Bjorken process is favoured over the  $WW$  fusion because of the absence of missing energy. The reach also depends on the various choices of the beam polarization options. To compare with the possibility of unpolarized beams, the reach of  $\mathcal{F}$  is  $2.1 \times 10^{-4}$  for identical  $\sqrt{s}$  and integrated luminosity, when the polarization option is switched off.

From the above discussion as well as from the tables, it is obvious that the discovery will most probably be made at the low- $\sqrt{s}$  options. This is easy to understand: the Bjorken cross-section falls with increasing  $\sqrt{s}$ , which is nothing but a propagator suppression of the amplitude. At  $\sqrt{s} = 1 \text{ TeV}$ , the VBF cross-section becomes important. The  $WW$  fusion has neutrinos associated with it and therefore one must take care of the background coming from  $e^+e^- \rightarrow W^+W^-$  followed by leptonic decays of the  $W$ .

The job of the theoreticians is to translate the model parameters, for both LH and DH types to  $\mathcal{F}$ . If the new Yukawa couplings are predicted by the model, the experiment will give the mixing angle  $\theta$  between the SM doublet and the new scalar. For the DH models, the data will give the shift from the alignment limit.

The ILC will similarly be a good probe for any LFV couplings of the Higgs, like  $h \rightarrow e\mu$  or  $h \rightarrow e\tau$ . The reaches will be quite similar as one looks at the Bjorken process followed by hadronic  $Z$  decays. Note that the  $e\mu$  final state will have a detection efficiency better than final states containing a  $\tau$  lepton, and so the reach should improve. However, there is no hint for any excess in these channels as yet [15], and so we do not go into any discussion of them.

*Note added:* When the paper was complete, Ref. [21] came to the arXiv which also treats the LFV decays at the ILC, also including the possible 1 TeV option.

### Acknowledgements

I.C. acknowledges the Council for Scientific and Industrial Research, Government of India, for a research fellowship. A.K. acknowledges the Department of Science and Technology, Government of India, and the Council for Scientific and Industrial Research, Government of India, for support through research grants.

## References

- [1] V. Khachatryan *et al.* [CMS Collaboration], Phys. Lett. B **749**, 337 (2015) [arXiv:1502.07400 [hep-ex]].
- [2] G. Aad *et al.* [ATLAS Collaboration], arXiv:1604.07730 [hep-ex];  
G. Aad *et al.* [ATLAS Collaboration], arXiv:1508.03372 [hep-ex].
- [3] B. Bhattacharjee, S. Chakraborty and S. Mukherjee, arXiv:1505.02688 [hep-ph].
- [4] A. Crivellin, G. D'Ambrosio and J. Heeck, Phys. Rev. Lett. **114**, 151801 (2015) [arXiv:1501.00993 [hep-ph]].
- [5] D. Das and A. Kundu, Phys. Rev. D **92**, no. 1, 015009 (2015) [arXiv:1504.01125 [hep-ph]].
- [6] D. Aristizabal Sierra and A. Vicente, Phys. Rev. D **90**, no. 11, 115004 (2014) [arXiv:1409.7690 [hep-ph]];  
I. Dorsner, S. Fajfer, A. Greljo, J. F. Kamenik, N. Kosnik and I. Nisandzic, JHEP **1506**, 108 (2015) [arXiv:1502.07784 [hep-ph]];  
I. de Medeiros Varzielas, O. Fischer and V. Maurer, JHEP **1508**, 080 (2015) [arXiv:1504.03955 [hep-ph]].



- W. Altmannshofer, S. Gori, A. L. Kagan, L. Silvestrini and J. Zupan, Phys. Rev. D **93**, no. 3, 031301 (2016) [arXiv:1507.07927 [hep-ph]].
- [7] M. Sher and K. Thrasher, Phys. Rev. D **93**, no. 5, 055021 (2016) [arXiv:1601.03973 [hep-ph]].
- [8] A. Pilaftsis, Phys. Lett. B **285**, 68 (1992);  
C. Alvarado, R. M. Capdevilla, A. Delgado and A. Martin, arXiv:1602.08506 [hep-ph].
- [9] V. Khachatryan *et al.* [CMS Collaboration], Eur. Phys. J. C **75**, no. 5, 212 (2015) [arXiv:1412.8662 [hep-ex]];  
M. Flechl [CMS Collaboration], arXiv:1510.01924 [hep-ex].
- [10] V. Khachatryan *et al.* [CMS Collaboration], Phys. Lett. B **736**, 64 (2014) [arXiv:1405.3455 [hep-ex]].
- [11] G. Moortgat-Pick *et al.*, Eur. Phys. J. C **75**, no. 8, 371 (2015) [arXiv:1504.01726 [hep-ph]].
- [12] S. Kanemura, K. Matsuda, T. Ota, T. Shindou, E. Takasugi and K. Tsumura, Phys. Lett. B **599**, 83 (2004) [hep-ph/0406316].
- [13] F. J. Botella, G. C. Branco, M. Nebot and M. N. Rebelo, arXiv:1508.05101 [hep-ph].
- [14] D. Chowdhury and O. Eberhardt, JHEP **1511**, 052 (2015) [arXiv:1503.08216 [hep-ph]];  
J. Bernon, J. F. Gunion, H. E. Haber, Y. Jiang and S. Kraml, Phys. Rev. D **92**, no. 7, 075004 (2015) [arXiv:1507.00933 [hep-ph]].
- [15] CMS Collaboration [CMS Collaboration], CMS-PAS-HIG-14-040.
- [16] T. P. Cheng and M. Sher, Phys. Rev. D **35**, 3484 (1987);  
J. L. Diaz-Cruz and J. J. Toscano, Phys. Rev. D **62**, 116005 (2000) [hep-ph/9910233];  
T. Han and D. Marfatia, Phys. Rev. Lett. **86**, 1442 (2001) [hep-ph/0008141];  
J. I. Aranda, F. Ramirez-Zavaleta, J. J. Toscano and E. S. Tututi, Phys. Rev. D **78**, 017302 (2008) [arXiv:0804.2652 [hep-ph]];  
A. Fernandez, C. Pagliarone, F. Ramirez-Zavaleta and J. J. Toscano, J. Phys. G **37**, 085007 (2010) [arXiv:0911.4995 [hep-ph]].
- [17] S. i. Kawada, K. Fujii, T. Suehara, T. Takahashi and T. Tanabe, Eur. Phys. J. C **75**, no. 12, 617 (2015) [arXiv:1509.01885 [hep-ex]].
- [18] G. Aarons *et al.* [ILC Collaboration], arXiv:0709.1893 [hep-ph],  
J. Brau *et al.* [ILC Collaboration], arXiv:0712.1950 [physics.acc-ph].
- [19] T. Abe *et al.* [Linear Collider ILD Concept Group - Collaboration], arXiv:1006.3396 [hep-ex].
- [20] A. Belyaev, N. D. Christensen and A. Pukhov, Comput. Phys. Commun. **184**, 1729 (2013) [arXiv:1207.6082 [hep-ph]].
- [21] S. Banerjee, B. Bhattacharjee, M. Mitra and M. Spannowsky, arXiv:1603.05952 [hep-ph].